HITACHI

Implementation Guide

Configuring Global-Active Device and VMware vSphere Metro Storage Cluster on a Hitachi Storage System

MK-SL-415-05
Hitachi Vantara

Table of Contents

Notices and Disclaimer

About This Document

Introduction

This document provides guidelines for configuring Global-active device (GAD) and VMware vSphere Metro Storage Cluster (vMSC) using a Hitachi storage system.

Document Revisions

Revision Number	Date	Details
v1.0	August 2020	Initial release
v2.0	October 2022	iSCSI host connection support removal for GAD
v3.0	October 2023	ESXi8.0 support addition
v4.0	January 2025	Hitachi VSP One B20 support addition, ESXi 6.5 and 6.7 support removal
v5.0	November 2025	Hitachi VSP One Block Series addition VCF 9.x support addition ESXi 7.0 support removal

Accessing Product Downloads

Product software, drivers, and firmware downloads are available on Hitachi Vantara Support Connect: https://support.hitachivantara.com/.

Log in and select Product Downloads to access the most current downloads, including updates that may have been made after the release of the product.

Getting Help

Hitachi Vantara Support Connect is the destination for technical support of products and solutions sold by Hitachi Vantara. To contact technical support, log on to Hitachi Vantara Support Connect for contact information: https://support.hitachivantara.com/en_us/contact-us.html.

Hitachi Vantara Community is a global online community for customers, partners, independent software vendors, employees, and prospects. It is the destination to get answers, discover insights, and make connections. **Join the conversation today!** Go to community.hitachivantara.com, register, and complete your profile

Comments

Send any comments on this document to Docs-Feedback@hitachivantara.com. Include the document title, including the revision level, and refer to specific sections and paragraphs whenever possible. All comments become the property of Hitachi Vantara LLC.

Thank you.

Introduction

A VMware vSphere Metro Storage Cluster (vMSC) is a specific storage configuration that combines replication with array-based clustering. These solutions are typically deployed in environments such as metropolitan or campus, where the distance between data centers is limited.

Purpose

This document provides instructions for configuring a VMware vMSC across two data centers using Hitachi storage systems. Additionally, it includes various failure scenarios based on the use case.

vMSC Block Diagram

The following diagram shows the metro cluster between two datacenters and the quorum storage placed at a third site:

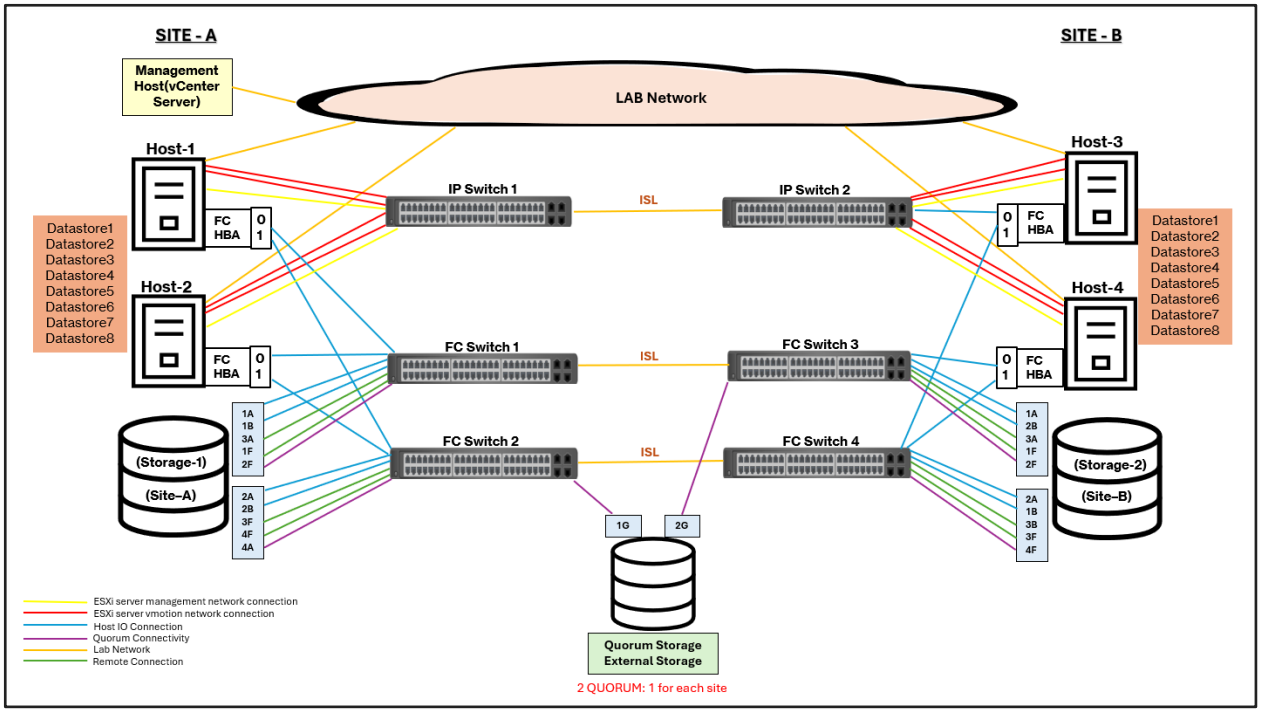


Figure 1: vMSC block diagram

Hardware Requirements

The hardware used for vMSC tests must be listed as supported in the VMware Compatibility Guide (VCG). The following components are required to create a VMware vMSC environment:

- Four ESXi 8.0 or VCF 9.x hosts with two in each site, having four port NIC card and two port FC HBA as per requirement. All hosts will be running two VMs each (minimum).
- Two IP switches for management network and vMotion connections between hosts/VMs.
- Four FC switches for SAN connectivity to the datacenter storage network.
- Two GAD Quorum Disks (iSCSI disk from virtual machine or separate FC/iSCSI storage system such as the Hitachi VSP F/G or other supported third-party storage systems).

  One management host (vCenter server). For granular level support information on protocols for host connectivity, remote connectivity and VSP SVOS levels, check product compatibility guide https://compatibility.hitachivantara.com/products/storage-certification?partner=19&certification=92

Configuring vMSC with GAD

Configuring vMSC with GAD consists of the following high-level steps:

1. Host configuration
2. Switch configuration
3. Primary and secondary storage systems configuration
4. Quorum storage configuration
5. GAD configuration

Host configuration

The following components are required for configuring the host:

- For vSphere 8.0 ESXi or VCF 9.x host servers, the physical hosts running the virtual machines in both data centers must be managed by the vCenter Server.
- For connecting SAN to the datacenter storage network, you must connect the Host HBAs to FC switches.
- For configuring vMotion between hosts, you must connect one NIC port of a two port NIC card on each host to IP switches of same site.
  - Create a Standard Virtual Switch (vSwitch) by adding the physical uplink that is connected to the IP switch.
  - Create VMkernel network adapters by assigning static IP addresses and enable vMotion service on each host to complete vMotion configuration.
- ESXi hosts using the iSCSI protocol with GAD is not supported by any VSP storage system.
- All four hosts along with the VMs must be a part of a single vSphere Cluster under a VMware Datacenter.
- For all GAD configurations, ATS (Atomic Test and Set) heartbeat must be disabled on all ESXi hosts. To disable ATS, run the following command on all ESXi hosts.

```
esxcli system settings advanced set -i 0 -o /VMFS3/UseATSForHBOnVMFS5
```

- Two VMs are located in each host. RDM LUNs or Virtual Machine File System (VMFS) volumes from the datastore must be assigned to those VMs.
- Host multipathing: You must configure all four hosts with multipathing software, either with VMware Native Multipathing (NMP) or with Hitachi Dynamic Link Manager (HDLM) to load-balance I/O between all available preferred and non-preferred paths.
  - NMP or HDLM multipathing software integrates with GAD to provide load balancing, path optimization, path failover, and path failback capabilities for vSphere hosts.
- When HDLM is selected as the multipathing software, the HDLM for VMware zip file is copied to the required server. Then, after unzipping the file, five vib files appear under five different folders. You must install all vib files first to install HDLM on the server as follows:

```
esxcli software vib install -v /<HDLM location>/vib20/hex-hdlm-dlnkmgr/*.vib

esxcli software vib install -v /<HDLM location>/vib20/psp-hdlm-exlbk/*.vib

esxcli software vib install -v /<HDLM location>/vib20/psp-hdlm-exlio/*.vib

esxcli software vib install -v /<HDLM location>/vib20/psp-hdlm-exrr/*.vib

esxcli software vib install -v /<HDLM location>/vib20/satp-hdlm/*.vib
```

- To verify whether Hitachi LUNs are managed by HDLM, run the following command:

```
esxcli storage nmp device list
```

The following shows the output for a LUN:

```
Device Display Name: HITACHI Fibre Channel Disk (naa.60060e8008753e000050753e00000133)
```



```
Storage Array Type: HTI_SATP_HDLM

Storage Array Type Device Config: {device config options }

Path Selection Policy: HTI_PSP_HDLM_EXLIO

Path Selection Policy Device Config:

Path Selection Policy Device Custom Config:

Working Paths: vmhba65:C0:T0:L0, vmhba64:C0:T0:L0

Is USB: false
```

Switch configuration

The following components are required for configuring the switch:

- For connecting LAN to the datacenter network from each host and VM, two network switches must be configured.
- For connecting hosts of both sites to the datacenter storage network (FC), four FC switches must be configured.
- For configuring the metro cluster environment, FC switch 1 in site-1 must be in cascade with FC switch-3 in site-2, and similarly, FC switch 2 in site-1 must be in cascade with FC switch-4 in site-2.
- For configuring vMotion and network management, IP switch 1 in site-1 must be in cascade with IP switch 2 in site-2.

Primary and secondary storage systems configuration

The following components are required for configuring the primary and secondary storage system:

- Two Hitachi storage systems configured (one in each site) and connected to FC/iSCSI switches of the respective site.
- Remote connectivity for GAD is supported for both the FC and iSCSI protocol.
- Each site storage must have two pairs of LUNs, one set for primary volumes and another set for secondary volumes. These LUNs must be provisioned to all the four hosts on both sites.
- Each site storage must have two pairs of storage ports for MCU (initiator) and RCU (target) pair. These connections act as the storage replication link between the primary and secondary storage systems.
- Site-1 storage primary volumes are in GAD pair with secondary volumes of the site-2 storage system. Similarly, site-2 storage primary volumes are in GAD pair with secondary volumes of the site-1 storage system.
- For NMP-ALUA configurations, the ALUA setting must be enabled on the P-VOLs for both sites. For enabling ALUA on Hitachi LUNs, run the following command:

  ```
  raidcom modify ldev -ldev_id <ldev_id> -alua enable -fx -IH<horcm_instance>
  ```

  For example:

  ```
  raidcom modify ldev -ldev_id 08:10 -alua enable -fx -IH4545
  ```

- Path optimization settings must be configured on the primary and secondary storage host groups as follows:

  ```
  raidcom modify lun -port cl1-d HOSTGROUP -lun_id all -asymmetric_access_state optimized -
    I10 (On PVOL host group)

  raidcom modify lun -port cl1-d HOSTGROUP -lun_id all -asymmetric_access_state non_optimized
    -I10 (On SVOL host group)
  ```

- HMO78 must be set on the Host group having S-VOLs for all HDLM configurations. This is not required for NMP-ALUA configurations.

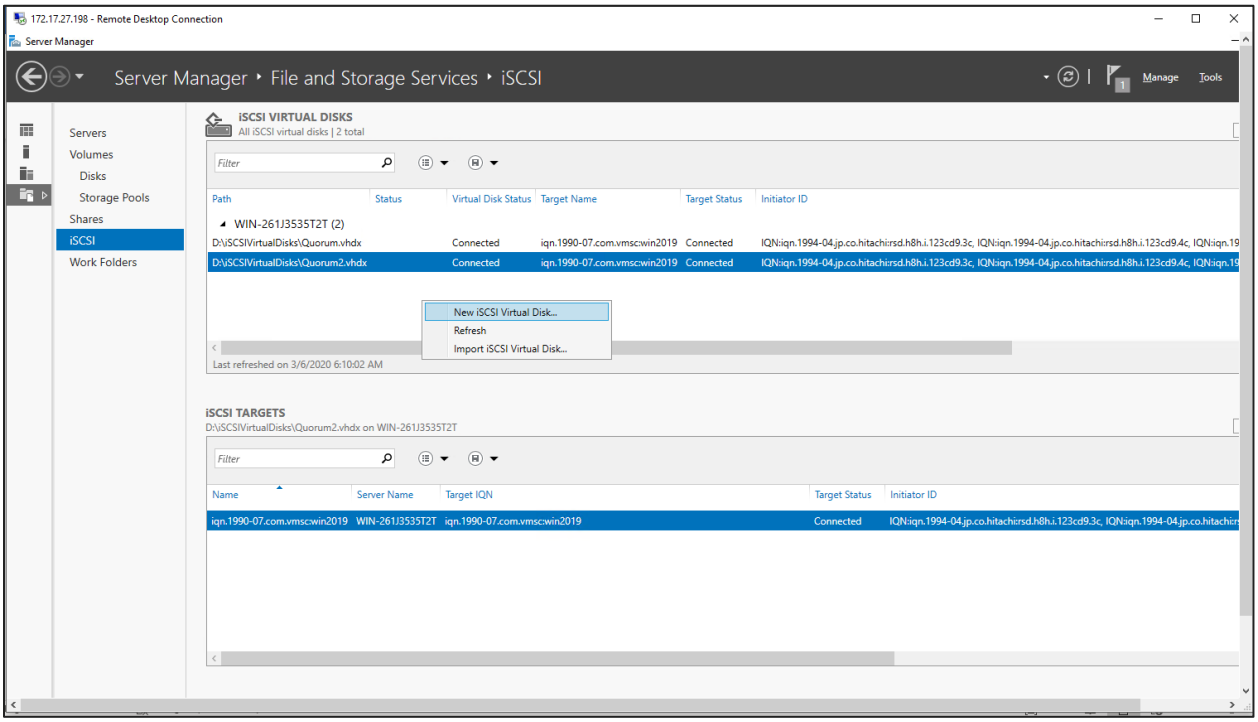
Quorum storage configuration

The following components are required for configuring the Quorum storage:

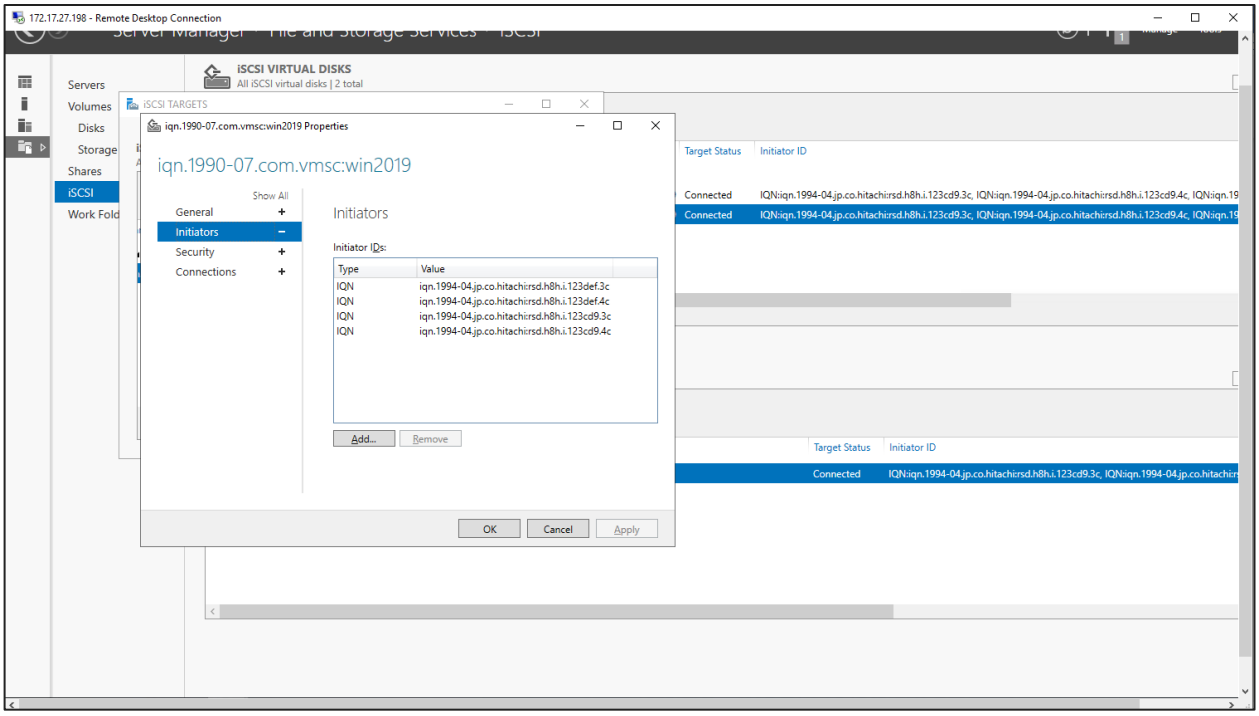
- Two quorum disks assigned for each site GAD pair set.
- The Quorum disk can be configured by either assigning an iSCSI disk from the local disk of the server hosting a Microsoft Windows Server or a separate storage system such as the Hitachi VSP F/G or other supported third-party storage systems.

To configure a Windows Server local disk as an iSCSI Quorum disk for VSP storage systems, complete the following steps:

1. Navigate to **Server Manager > File and Storage Services > iSCSI > New iSCSI Virtual Disk**.



2. Under iSCSI Targets, right-click **View all Targets** > **Properties** > **Initiators** and add the IQN of the storage ports to use as External Ports for the Quorum.



GAD configuration

Two sets of GAD pairs configured using the storage UI or CCI for each site storage for the vMSC environment are required for GAD configuration.

To create a GAD pair from Raid Manager CCI server on the primary storage system, complete the following steps:

1. Create LDEV for the DP pool by running following command:

   ```
   raidcom add ldev -parity_grp_id <> -ldev_id <ldev_id> -capacity <ldev capacity> -
   I<primary_storage_horcm_instance>
   ```

2. Format the newly created LDEV by running the following command:

   ```
   raidcom initialize ldev -ldev_id <ldev_id> -operation fmt -
   I<primary_storage_horcm_instance >
   ```

3. Create a DP pool by running the following command for primary volumes in the site-1 storage system. Repeat the procedure to create a DP pool for primary volumes in the site-2 storage system.

   ```
   raidcom add dp_pool -pool_name <pool_name> -ldev_id <ldev_id> -I<Primary storage horcm
   instance>
   ```

4. Create LDEVs from this DP pool for primary volumes in the site-1 storage system by running the following command. Repeat the procedure to create primary volumes in the site-2 storage system.

   ```
   raidcom add ldev -pool <pool_id> -ldev_id <ldev_id> -capacity <pool_capacity> -
   I<primary_storage_horcm_instance>
   ```

   From SVOS 10.4.1 or later for Hitachi VSP One Block models, only DRS volume is supported. To create a DRS volume, run the following command:

   ```
   raidcom add ldev -pool <pool_id> -ldev_id <ldev_id> -capacity <capacity> -capacity_saving
   deduplication_compression -drs -request_id auto -<horcm_instance>
   ```

5. Format the newly created LDEV by running the following command:

   ```
   raidcom initialize ldev -ldev_id <ldev_id> -operation fmt -
   I<primary_storage_horcm_instance >
   ```


6. Create a host group and set host mode options for primary volumes host group in the site-1 storage system. Repeat the procedure to create primary volumes host group in the site-2 storage system.
7. Create a host group by running the following command:

   ```
   raidcom add host_grp -port <host_group_id> -host_grp_name <host_group_name> -
   I<primary_storage_horcm_instance>
   ```

8. Set the host mode and host mode options by running the following command:

   ```
   raidcom modify host_grp -port <host_group_id> -host_mode 21 -host_mode_opt 54 63 114  -
   I<primary_storage_horcm_instance>
   ```

9. Set the Port topology and add HBA port wwns (for all four hosts) at Host Group by running the following commands:

   ```
   raidcom modify port -port <port_id>  -port_speed 0 -topology f_port -security_switch  y -
   I<primary_storage_horcm_instance>
   ```

   ```
   raidcom add hba_wwn -port <host_group_id> -hba_wwn <HBA_WWN> -
   I<primary_storage_horcm_instance>
   ```

10. Assign LDEVs to the host group by running the following command:

    ```
    raidcom add lun -port <host_group_id> -lun_id 0 -ldev_id <ldev_id> -
    I<primary_storage_horcm_instance>
    ```

To create a GAD pair from Raid Manager CCI server on the secondary storage system, complete the following steps:

1. Create a resource group of the primary storage system (site-1) on the secondary storage system (site-2) and assign the respective secondary resources. Repeat the procedure to create a resource group of the primary storage system (site-2) on the secondary storage system (site-1) and assign the respective secondary resources. Run the following command:

   ```
   raidcom add resource -resource_name <resource_group_name> -virtual_type
   <primary_storage_serial_number> <storage_model_type> -IH<secondary_storage_horcm_instance>
   ```

2. Reserve the host group ID in the resource group of the storage system at the secondary site by running the following command:

   ```
   raidcom add resource -resource_name <resource_group_name> -port <secondary_hostgroup_ID> -
   IH<secondary_storage_horcm_instance>
   ```

3. Delete the virtual LDEV ID of the volumes from the secondary storage system for creating GAD pairs by running the following command:

   ```
   raidcom unmap resource -ldev_id <LDEV_ID> -virtual_ldev_id <virtual_LDEV_ID> -
   IH<secondary_storage_horcm_instance>
   ```

4. Reserve the LDEV IDs in the resource group by running the following command:

   ```
   raidcom add resource -resource_name <resource_group_name> -ldev_id <LDEV_ID> -
   I<secondary_storage_horcm_instance>
   ```

5. Set the reservation attribute for GAD to the LDEV IDs by running the following command:

   ```
   raidcom map resource -ldev_id <LDEV_ID> -virtual_ldev_id  reserve -
   IH<secondary_storage_horcm_instance>
   ```

6. For the LDEV ID where the reservation attribute was set, ffff or 65535 is displayed for VIR_LDEV (virtual LDEV ID)

   You can verify this by running the following command:

   ```
   raidcom get ldev -ldev_id <LDEV_ID> -fx -IH<secondary_storage_horcm_instance>
   ```


7. Create a host group of the GAD secondary site storage system and set Host Mode Options by running the following commands:

```
raidcom add host_grp -port <Host_group_ID> -host_grp_name <host_group_name> -
IH<secondary_storage_horcm_instance>
```

```
raidcom modify host_grp -port <Host_group_ID> -host_mode 21 -host_mode_opt 54 63 78 114  -
IH<secondary_storage_horcm_instance>
```

   Note: Ensure that HMO78 is enabled on the Host Group associated with the S-VOLs for all HDLM configurations.

8. Set Port topology and add HBA port wwn of all the four hosts of both sites at the Host Group by running the following command:

```
raidcom modify port -port <Port_ID> -port_speed 0 -topology f_port -security_switch  y -
IH<secondary_storage_horcm_instance>
```

```
raidcom add hba_wwn -port <Host_group_ID> -hba_wwn <HBA_WWN> -
IH<secondary_storage_horcm_instance>
```

9. Create DP pool and LDEVs for secondary volumes of site-1 GAD pair by running the following command. Repeat the procedure to create DP pool and LDEVs for secondary volumes of site-2 GAD pair.

```
raidcom add dp_pool -pool_name <Secondary_pool_name> -ldev_id <Pool_volume_LDEV_ID> -
I<secondary_storage_horcm_instance>
```

10. Create secondary volumes with the same capacity as the primary volumes by running the following command:

```
raidcom add ldev -pool <pool_id> -ldev_id <LDEV_ID> -capacity <LDEV_size> -
IH<secondary_storage_horcm_instance>
```

11. Add an LU path to the secondary volume by running the following command:

```
raidcom add lun -port <Host_grp_ID> -lun_id 0 -ldev_id <LDEV_ID> -
IH<secondary_storage_horcm_instance>
```

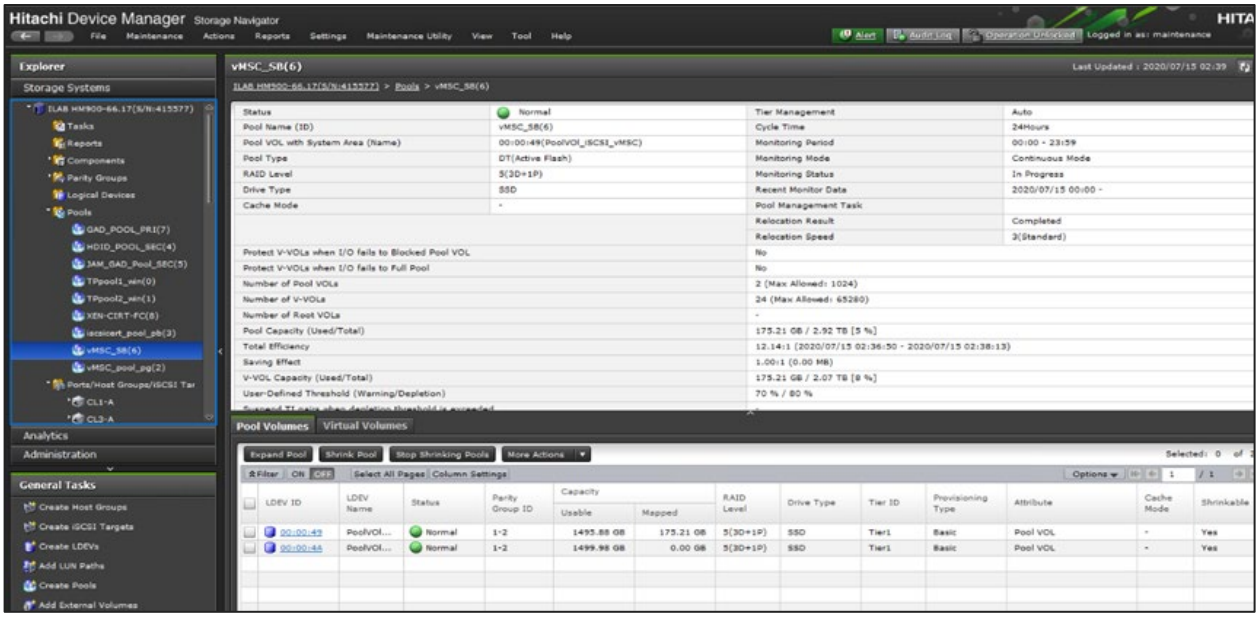
Creating a GAD Pair

To create a GAD Pair in site-1, run the following command in the site-1 storage system. Repeat the procedure to create site-2 GAD pair on the site-2 storage system.

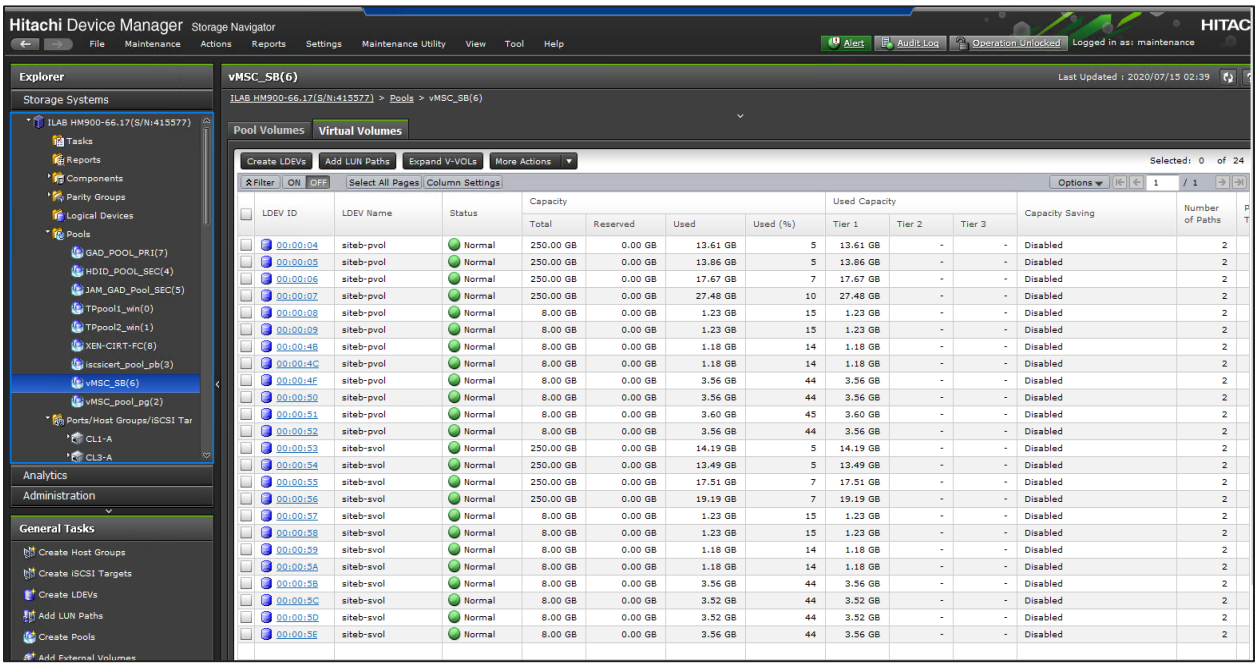
```
paircreate -g <GAD_PAIR_Name> -f never -vl -jq 7 -IH<primary_storage_horcm_instance>
```

To create GAD pair from SVP on the primary storage system, complete the following steps:

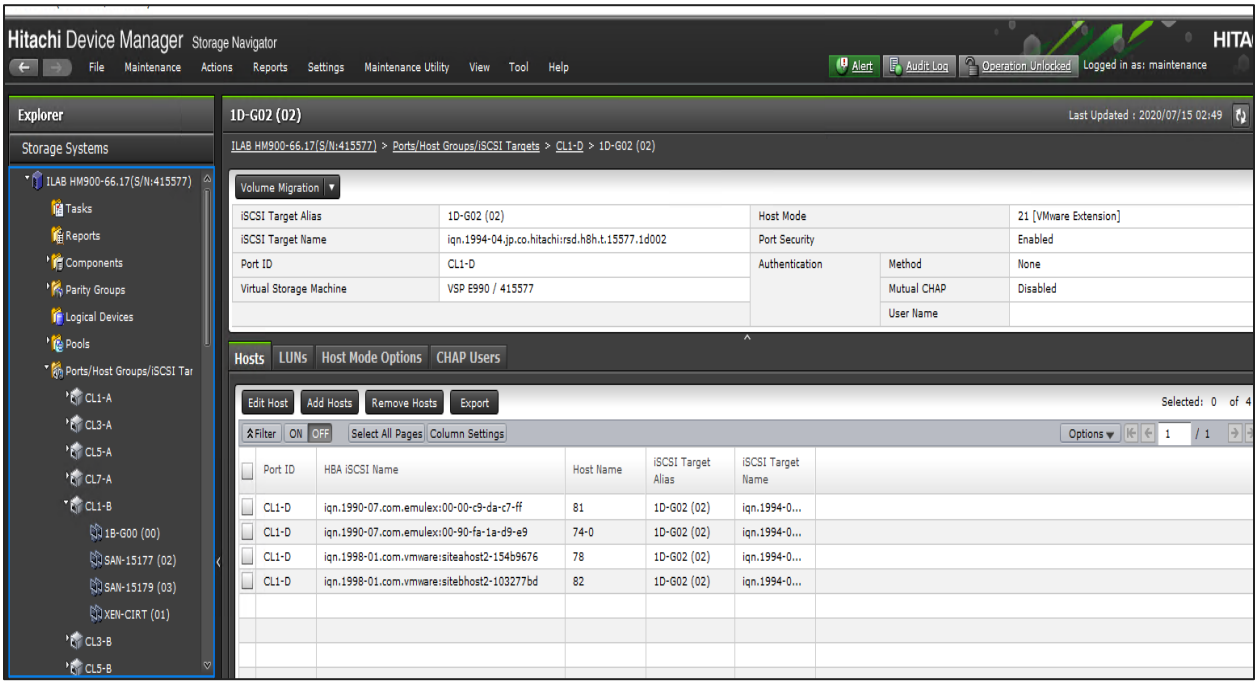
1. Create a DP pool for primary volumes on the site-1 storage system. Repeat the procedure to create a DP pool for primary volumes on the site-2 storage system.



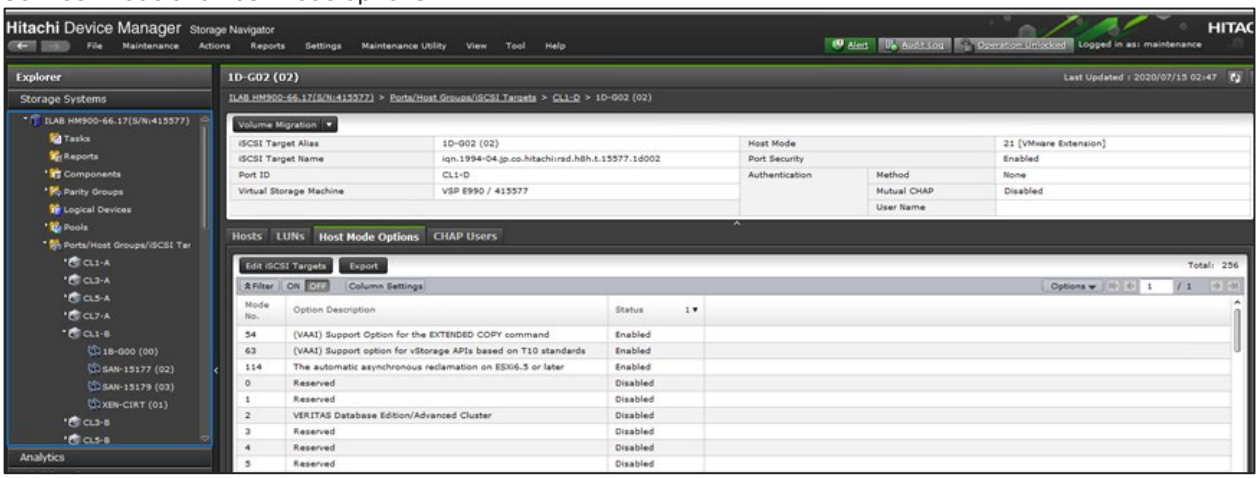
2. Create LDEVs from this DP pool for primary volumes in the site-1 storage system. Repeat the procedure to create primary volumes in the site-2 storage system.



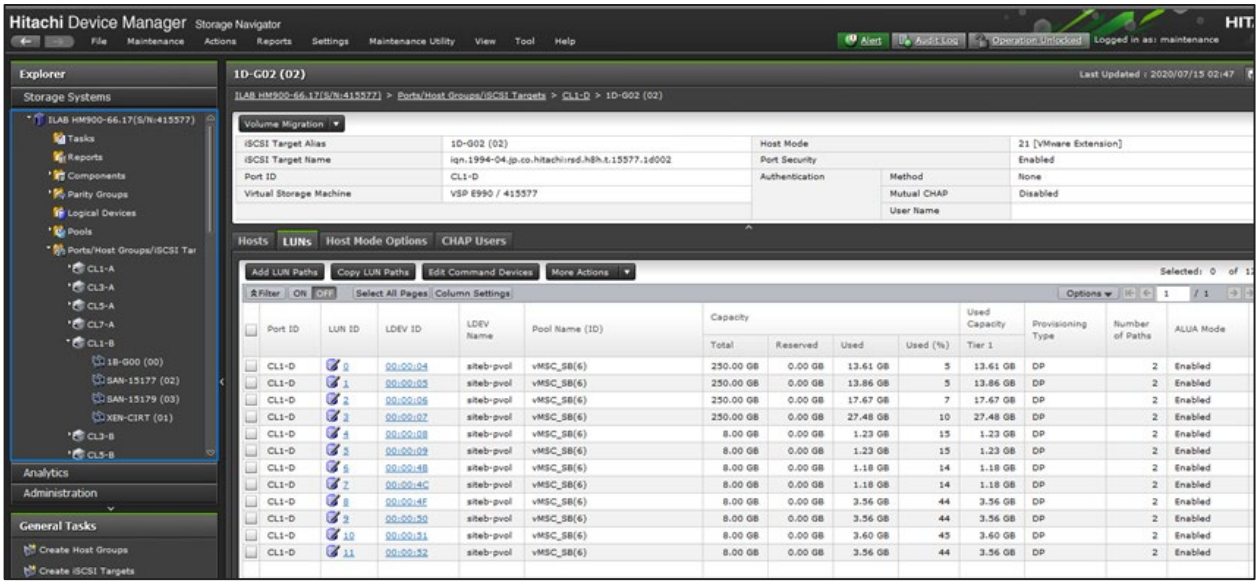
3. Create a host group and set the host mode options for primary volumes host group in the site-1 storage system. Repeat the procedure to create primary volumes host group in the site-2 storage system.



4. Set host mode and host mode options:

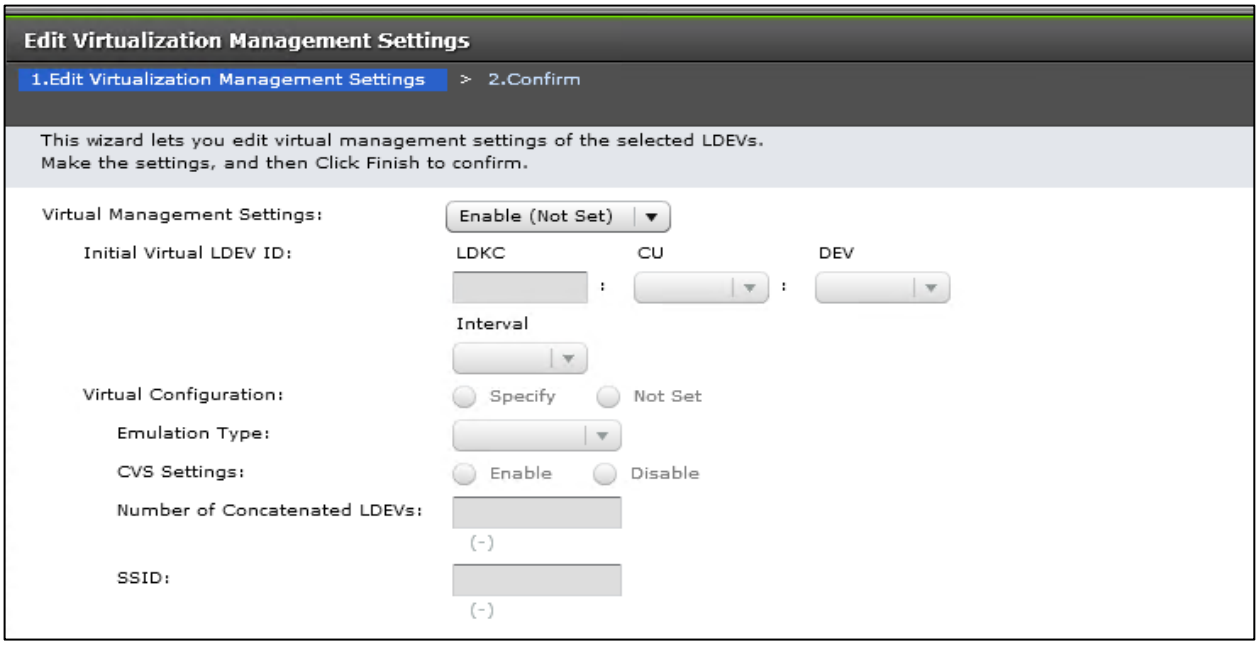


5. Assign LDEVs to the host group:



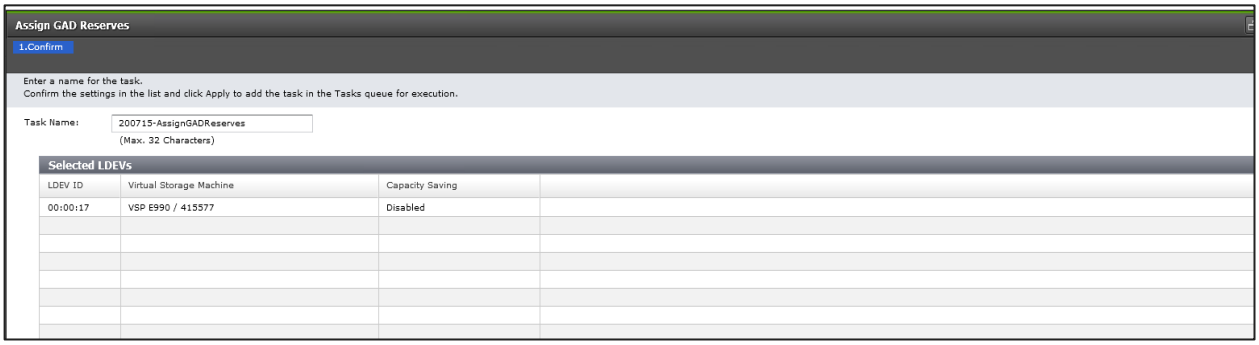
To create GAD pair from SVP on the secondary storage system, complete the following steps:

1. Create a resource group of the primary storage (site-1) type on the secondary storage system (site-2) and assign the respective secondary resources. Repeat the procedure to create a resource group of the primary storage (site-2) type on the secondary storage system (site-1) and assign the respective secondary resources.
2. Delete the virtual LDEV ID of the volumes from the secondary storage system for creating GAD pairs.

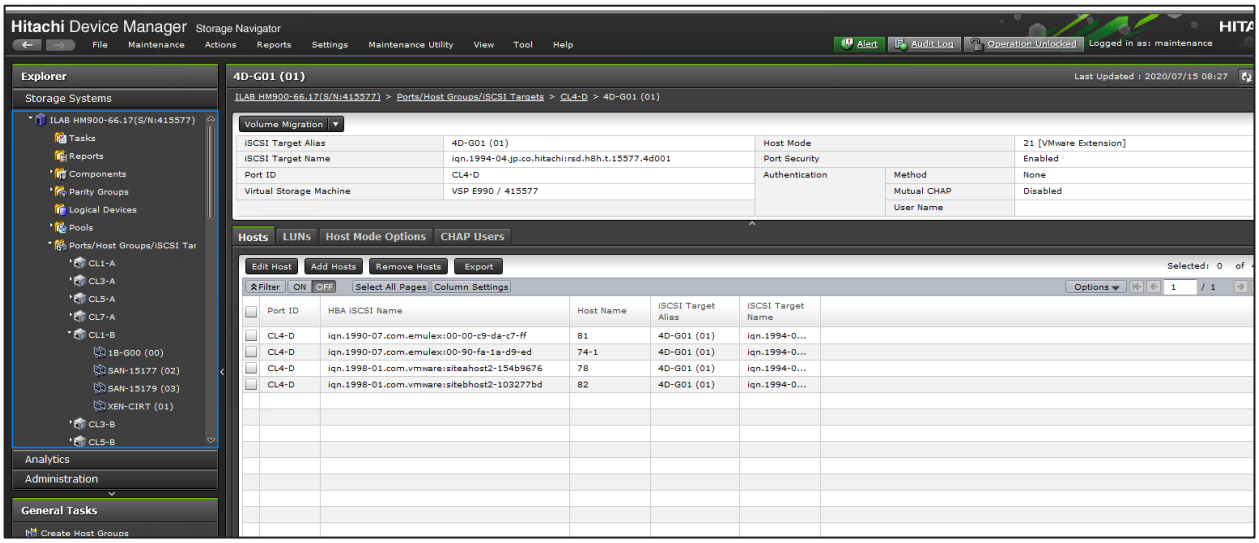


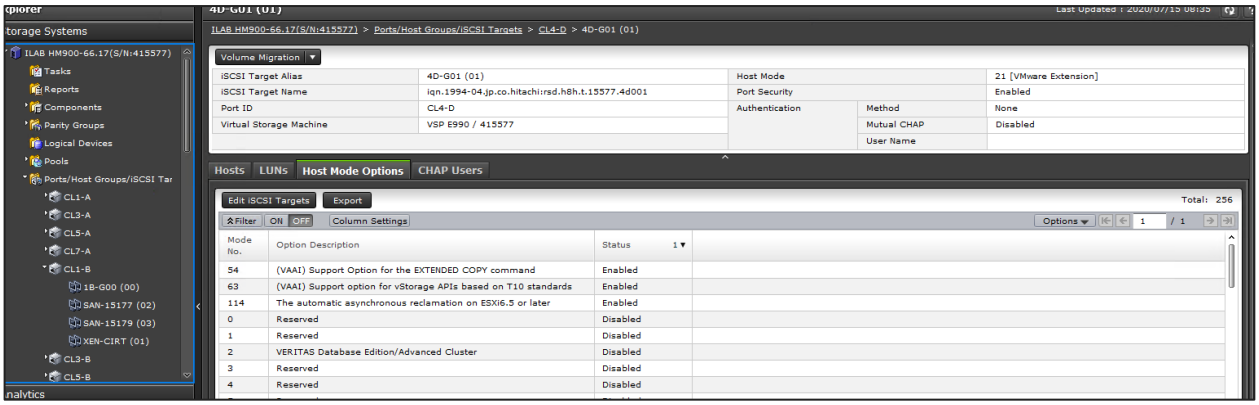
3. Set the reservation attribute to the volume for the secondary volume of the GAD pair.

4. Set the reservation attribute for GAD to the LDEV IDs.



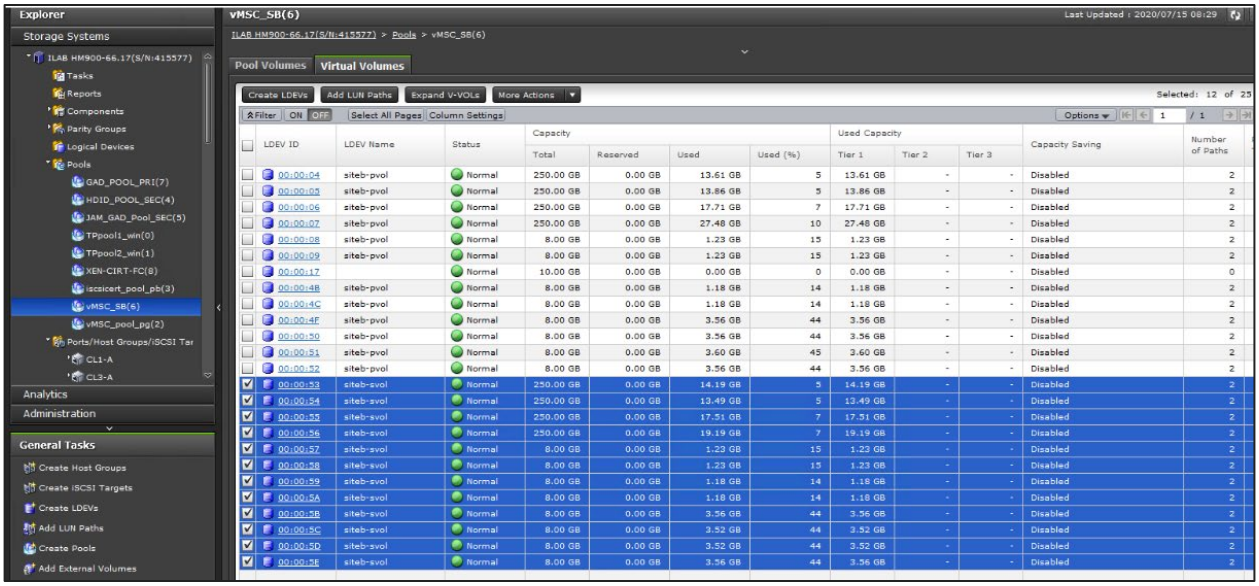
5. Create a host group of the GAD secondary site storage system and set Host Mode options.



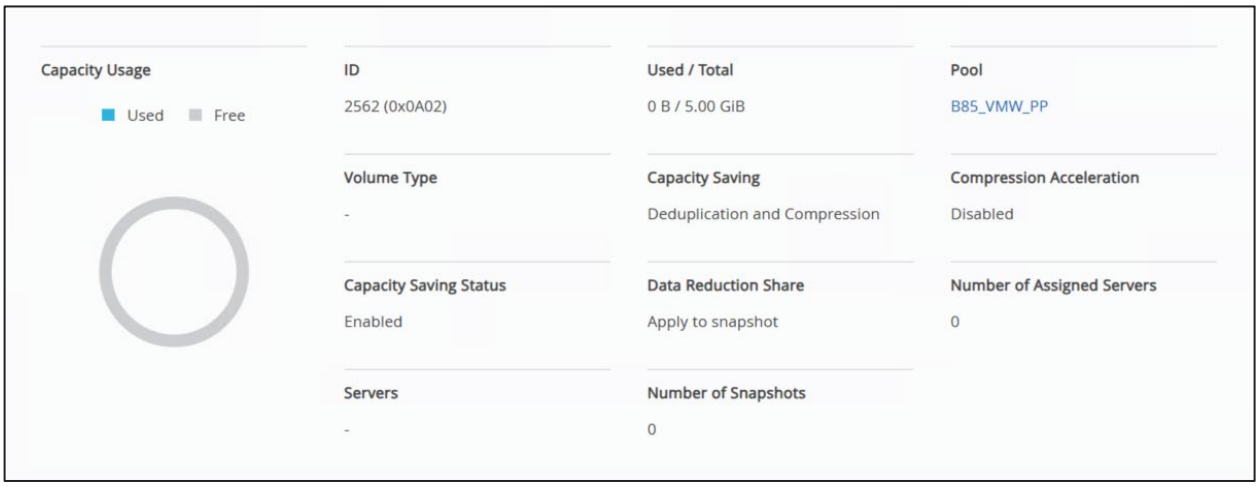


6. Create DP Pool and LDEVs for secondary volumes of site-1 GAD pair. Repeat the procedure for creating DP Pool and LDEVs for secondary volumes of site-2 GAD pair.

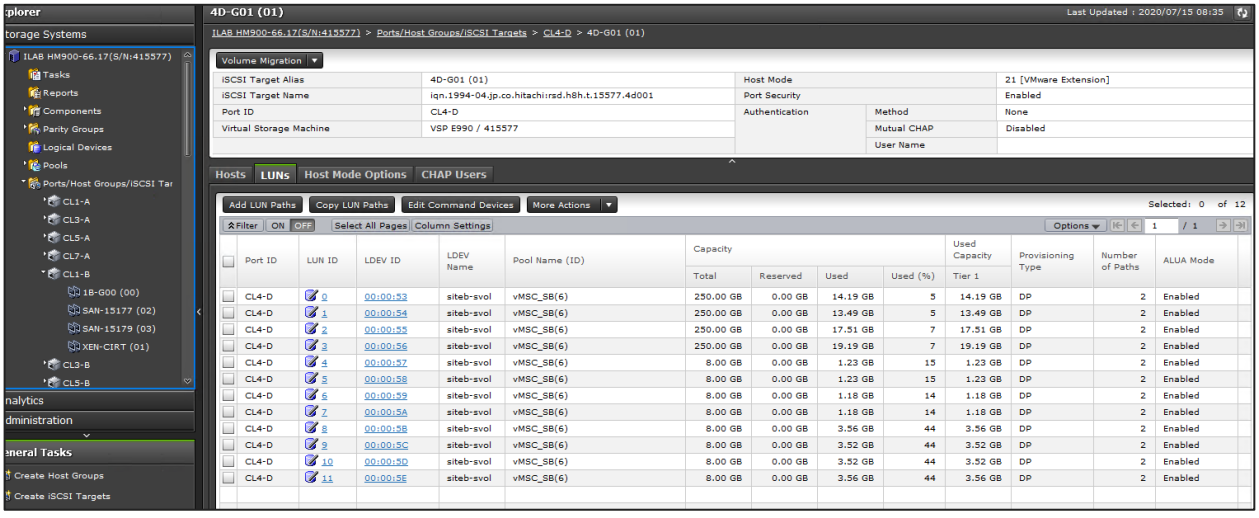
7. Create secondary volumes with the same capacity as the primary volumes.



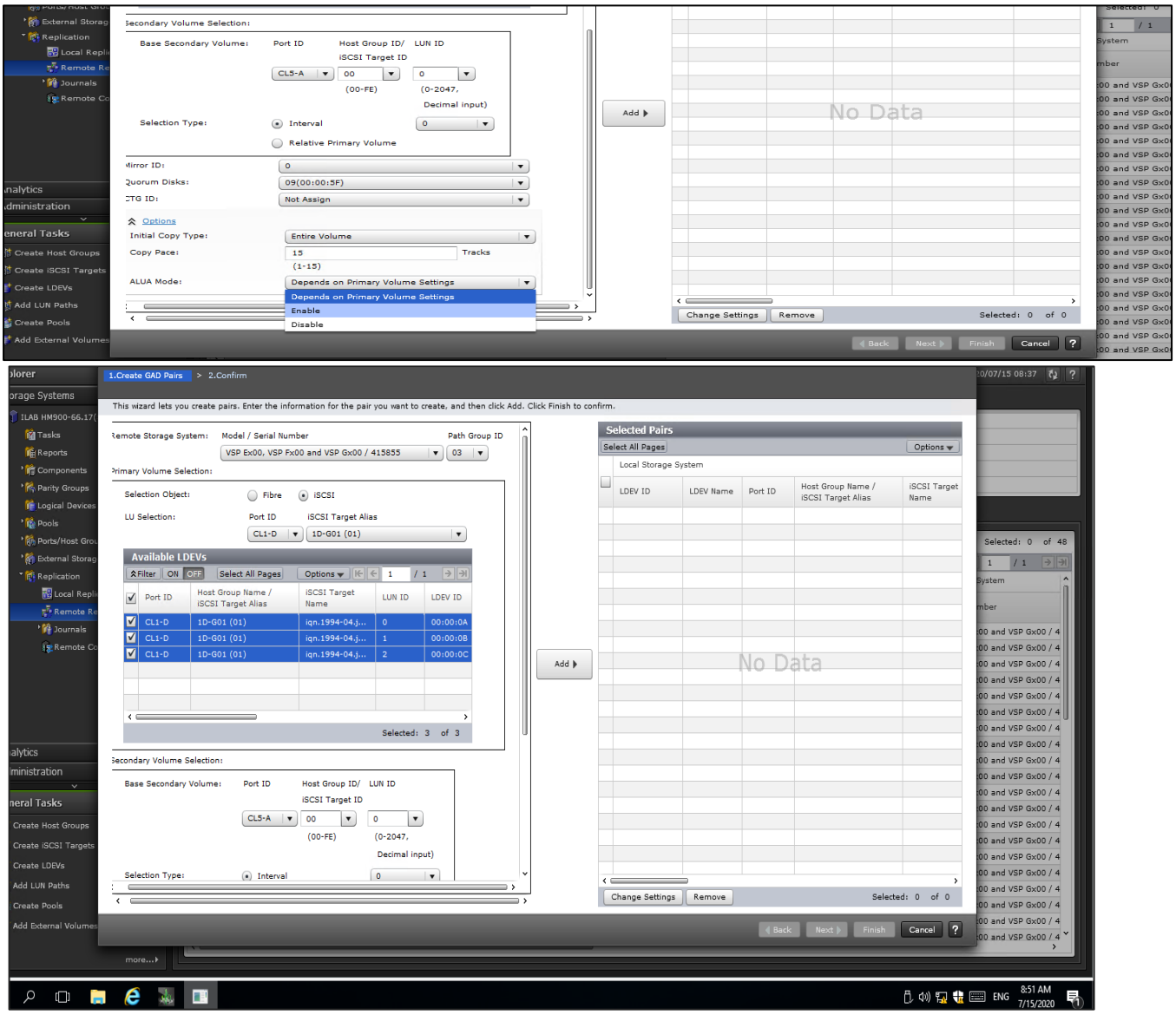
From SVOS 10.4.1 or later for Hitachi VSP One Block models, only DRS volume is supported.



8. Add an LU path to the secondary volume.



9. Create a GAD Pair.



For NMP-ALUA configuration, while creating an ALUA pair, the ALUA mode must be enabled.

Remote Connection Failure Scenarios

This section covers the typical failure scenarios in a GAD vMSC environment, along with the results for each scenario.

GAD pair behaves differently for failing remote connections for each site.

Test simulation results

Test simulation result of storage TC ports failure (ALUA/NMP) for a particular site (A or B):

Disabled site A storage TC ports:

- Site A P-VOLs win and site B corresponding S-VOLs block.
- Site B P-VOLs win and site A corresponding S-VOLs block.

Disabled site B storage TC ports:

- Site A P-VOLs block and site B corresponding S-VOLs win.
- Site B P-VOLs win and site A corresponding S-VOLs block.

Test simulation result of storage TC ports failure (HDLM) for a particular site (A or B):

Disabled site A storage TC ports:

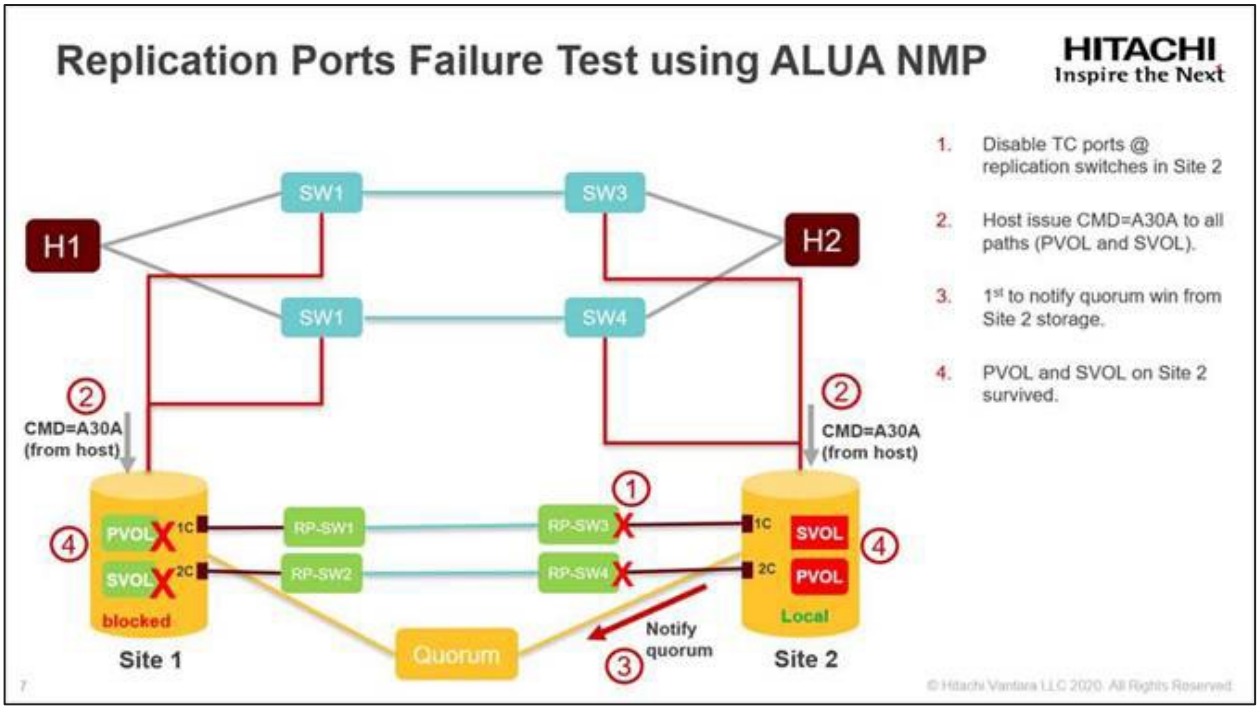
- Site A P-VOLs win and site B corresponding S-VOLs block.
- Site B P-VOLs win and site A corresponding S-VOLs block.

Disabled site B storage TC ports:

- Site A P-VOLs win and site B corresponding S-VOLs block.
- Site B P-VOLs win and site A corresponding S-VOLs block.

Note:

- For NMP/ALUA multipath, set HMO78=OFF.
- For NMP/ALUA, ensure that ALUA is enabled per LUN / Dedicated Ports for P-VOLs, and S-VOLs enabled with HG Optimized and Non-optimized Paths.
- For vSphere, the UI on all ESXi hosts shows the LUN status: “Active (IO)” à P-VOLs and “Active” à S-VOLs.
- For S-VOLs Storage Ports, zero IOPS was observed, and for P-VOLs storage ports, generated IO workload was observed.
- For NMP/ALUA, the host sends CMD=A30A to all the paths, and the storage that notifies the Quorum first survives.
- HDLM confirmed that no ALUA RTPG A3h command send occurred; therefore, both P-VOLs survived on both storage systems.


Replication Ports Failure Test using ALUA NMP
HITACHI
Inspire the Next
1. Disable TC ports @ replication switches in Site 2
2. Host issue CMD=A30A to all paths (PVOL and SVOL).
3. 1st to notify quorum win from Site 2 storage.
4. PVOL and SVOL on Site 2 survived.
H1
H2
SW1
SW3
SW1
SW4
CMD=A30A (from host)
CMD=A30A (from host)
PVOL
SVOL
1C
2C
blocked
Site 1
RP-SW1
RP-SW2
RP-SW3
RP-SW4
Quorum
Notify quorum
1C
2C
SVOL
PVOL
Local
Site 2
© Hitachi Vantara LLC 2020. All Rights Reserved.